OPERATIONAL RESEARCH IN EDUCATION

Jill Johnes
Department of Economics
Lancaster University
LA1 4YX
E: j.johnes@lancaster.ac.uk
T: +44 (0)1524 592102
F: +44 (0)1524 594244
December 2013
Revised October 2014

Abstract

Operational Research (OR) techniques have been applied, from the early stages of the discipline, to a wide variety of issues in education. At the government level, these include questions of what resources should be allocated to education as a whole and how these should be divided amongst the individual sectors of education and the institutions within the sectors. Another pertinent issue concerns the efficient operation of institutions, how to measure it, and whether resource allocation can be used to incentivise efficiency savings. Local governments, as well as being concerned with issues of resource allocation, may also need to make decisions regarding, for example, the creation and location of new institutions or closure of existing ones, as well as the day-to-day logistics of getting pupils to schools. Issues of concern for managers within schools and colleges include allocating the budgets, scheduling lessons and the assignment of students to courses. This survey provides an overview of the diverse problems faced by government, managers and consumers of education, and the OR techniques which have typically been applied in an effort to improve operations and provide solutions.

Keywords: Mathematical programming; Markov processes; Optimal control theory; Data envelopment analysis; Stochastic frontier analysis; Scheduling and timetabling;

Acknowledgements: I am grateful to my colleague Gerry Steele and to 2 anonymous referees for comments on an early version of the paper.
1. Introduction

Education covers a range of sectors from kindergarten, primary and secondary schooling, to post-compulsory and higher education. The expected years an individual might spend in education in total can vary considerably across countries; within the OECD, for example, a person in Indonesia can typically expect fewer than 14 years whilst one in Finland nearly 20 years (OECD 2013). An interesting dimension of education is that consumption at some levels is compulsory while at other levels it is voluntary; and because an individual’s consumption of education has both external and private benefits, it is often (but not exclusively) provided through public funding. This market failure and consequent public funding engender government intervention in the form of planning and resource allocation across the education sectors. These are complex areas, but they are ones where operational research (OR) tools can be effectively used to aid policy-makers (Platt 1962).

Top-level planning and resource allocation are not the only areas where OR can be useful. Education managers are faced with a plethora of problems in the day-to-day running of their institutions. These relate, for example, to the optimal allocation of their budget, or simply to where each class should take place and who should teach it. OR also has the tools to address these problems as is testified by the vast OR literature devoted to management, timetabling and scheduling in education.

OR originated as a tool to aid the military. In 1936, applied research into radar technology and its application in a military setting was undertaken jointly by British air force officers and civilian scientists. This led to the formation of operational research groups in the UK and operations research groups in the USA which brought together scientists from a variety of disciplines to solve problems encountered in a military context – encompassing the army, navy and air force (Gass 1994; Kirby 2003; Gass and Assad 2005; Weir and Thomas 2009). Once World War II was over, OR groups continued to be supported, with the focus switching to logistics, modelling and planning. It became apparent that OR had a place in solving operational problems in organisations unrelated to the military (Gass 1994), and so applications of OR techniques to business quickly followed the end the War. Indeed, the competitive advantage and consequent increase in profits enjoyed by firms which successfully applied new OR approaches in their operations were strong inducements to making OR an acceptable approach to solving problems in the business setting (Horvath 1955).

Operational researchers were, however, much slower to apply their skills in areas of public provision of services such as education, health, police and fire services. The lack of profit motive meant that there was a danger that these areas might remain ignored. Early publications called on operational researchers to become involved in studying the complex problems seen in provision of education (Horvath 1955; Platt 1962; Shepherd 1965; Dean 1968; Griffin 1968; Rath 1968) and demonstrated the relevance of OR tools in addressing these issues (Blaug 1967b; Van Dusseldorp et al. 1971). Education has been firmly on the OR agenda since that time.

This paper examines the following questions in the context of education. What types of problems has OR typically tried to address? Which OR tools are commonly applied? No attempt is made to provide a review of all OR applications to education but rather to give a flavour of the areas where OR tools have been used. While references are largely confined to mainstream OR journals, there is

---

1 Public spending on education averages 13% of total public expenditure across all OECD countries, and is more than 20% in some countries (OECD 2013).
inevitable reference to similar applications appearing in mainstream economics or education journals where there is often a parallel literature. Each section of this paper addresses a specific topic within the field of education and discusses the techniques which have been used to address it. The main areas of coverage are planning models (section 2); efficiency and performance (section 3); and routing and scheduling (section 4). Section 5 concludes and considers areas in education still to be explored by operational researchers.

2. Planning and resource allocation

The call for OR to be applied within education coincided with a burgeoning demand for education and training: the post-War period saw increasing birth rates in many Western economies as well as rapid economic and social changes which required an increasing supply of educated manpower (Blaug 1967a). There was a growing recognition that economic progress and growth required investment in both physical and human capital (Weisbrod 1962; Armitage et al. 1969). The expansion of education provision required accompanying resources, and so it was important to be able to predict student numbers at different education levels and hence resource requirements. Early forays by operational researchers into the field of education were therefore attempts at assisting education and manpower planners (see Schroeder 1973 for an early review).

2.1 Planning

The education sector can be seen as a series of components (i.e. different levels of education such as primary, secondary, vocational and tertiary) which are interconnected in such a way that each individual can follow a pathway which meets his own educational and training aspirations (Tavares 1995). Education is therefore a system; adopting this view allows operational researchers to model the system using a variety of approaches and provide useful forecasts for managers, planners and policy makers.

Planners are interested in projections of students and of needs (in terms of teachers and equipment) at all levels of education. Goal programming can be used to determine optimal numbers of students (at macro- and micro-levels) as demonstrated by an early study of vocational education in Missouri (Atteberry 1979) and another on determining the optimal admissions policy for an individual institution (Lee and Moore 1974).

A more commonly-used approach to educational planning, however, presents the education system as a series or flow of mathematical relationships (Van Dusseldorp et al. 1971). Studies differ in the mathematical representation – a simple Leontief input-output depiction of interdependence between students at various education levels (Stone 1965; 1966; Oliver and Hopkins 1972); a sequence of discrete events in time (students in different modules on a programme, for example) to which simulation can be applied (Saltzman and Roeder 2012); a Markov chain framework based on

---

2 The importance of OR in developing models for assisting in educational planning in the UK is revealed in Ladley (1987) who describes the models developed by the OR Unit of the Department of Education and Science.
students in each state (education level, for example) and their probability of moving to another state – but all are capable of providing forecasts of student numbers.

It is particularly attractive to view the education system in the framework of a Markov process which is defined as an ordered series of states linked by a transition matrix composed of probabilities of moving from one state to another. So in a college setting we might consider students to be in any one of the following states: studying full time; studying part time; on a temporary leave of absence; successfully graduated; or withdrawn (Kwak et al. 1986). From past data it is then possible to estimate the values of the transition matrix and use these to make predictions of student numbers at any stage.

Many examples of the application of mathematical models to education planning exist at both national and institution level (Gani 1963; Clough and McReynolds 1966; Armitage and Smith 1967; Correa 1967; Forecasting Institute of the Swedish Central Bureau of Statistics 1967; Thonstad 1967; Harden and Tcheng 1971; Massy 1976; Hopkins and Massy 1977; Smith 1978; Gray 1980; Nicholls 1983; 1985; Kwak et al. 1986; Brandeau et al. 1987). Because they are based on student flows from state to state, Markov chain models have proved particularly useful at the faculty and programme levels in providing not just predictions of students but also additional insights into, for example, non-completion both in postgraduate programmes (Bessent and Bessent 1980; Nicholls 2007) and undergraduate programmes (Shah and Burke 1999), required deployment of supervisors in a doctoral programme (Nicholls 2009), and evaluation of the efficacy of early-retirement programmes for university faculty (Hopkins 1974). There are fewer examples of the application of simulation to students flows; one such study, however, has proved useful in evaluating the potential effects on students, in terms of their time to complete the programme and graduation rates, of changes in curriculum provision brought about by recent budget cuts (Saltzman and Roeder 2012).

These planning models rely heavily on underlying assumptions such as those relating to the transition rates, and these in turn are often based on historical data. For planning at a school level, for example, the transition proportions will need to be adjusted if, for example, there is a change in birth rates, migration, expansion of educational provision in the local area, or increase in residential building in the catchment area (Smith 1978). More satisfactory models can be derived by altering the transition proportions to reflect additional uncertainty (Armitage et al. 1969; Massy et al. 1981). Even so, the models are highly descriptive and do not provide any indication of how or why the numbers observed in the system emerge. Only insofar as the system continues to behave in the future as in the past will projections be accurate.

2.2 Resource allocation

These mathematical flow models generally used in planning fail to answer the question of what is the optimal policy for planners (Alper et al. 1967; Correa 1967) and this leads on to the issue of optimal allocation of resources. Governments need to know not just how many students to expect at each education level, but also how much money is required to fund the predicted numbers. An individual education authority or school must also allocate its resources to provide education in line with predicted demand. But for the education system as a whole, or for an organization within the

---

3 Of course, if the model is set up in terms of staff or financial resources, then forecasts of those variables can be derived.
system, potential conflicts between competing objectives must be reconciled. This leads us into multi-objective decision-making in which goal programming is a popular methodological approach, and indeed has been the method of choice in addressing issues of optimal resource allocation at a national level (Benard 1967; Cobacho et al. 2010).

At the level of the organization, approaches to the problem of resource allocation within the institution have developed over time. Early pioneering work typically used goal programming to derive optimal resource allocation over a single period in a single ‘unit’ of an institution. For example, Lee and Clayton (1972) used goal programming to allocate funds from the central university to the different faculties within the institution; Geoffrion et al. (1972) used a similar approach in a graduate school context to allocate not monetary resources but rather staff time amongst the various tasks of teaching, research and administration.

Subsequent studies built on the approach of Geoffrion et al. (1972) either by integrating preference functions4, relating, for example, to academic outcomes, and using a mathematical programming model (usually goal programming) to obtain optimal budget decisions (Hopkins et al. 1977; Wehrung et al. 1978; Hemaida and Hupfer 1994/5; Kwak and Changwon 1998; Fandel and Gal 2001), or by extending to greater numbers of time periods and/or more units within the institution (Goyal 1973; Schroeder 1974; Walters et al. 1976; Sinuany-Stern 1984; Diminnie and Kwak 1986; Soyibo and Lee 1986; Caballero et al. 2001). Later more sophisticated approaches introduced feedback into the optimization models and solved using optimal control theory (OCT). This approach can be applied at a national perspective to identify, for example, optimal numbers of students at different education levels and hence to address the question of how the government should allocate resources amongst education sectors (Ritzen and Winkler 1979; Hartl 1983). In contrast, the same approach can be used at the level of the individual to identify the optimal education decision over a person’s lifetime (Southwick Jr and Zionts 1974; Pantelous and Kalogeropoulos 2009).

The difficulty with introducing complexity into such models is that the resulting problem can prove intractable; any simplification which can lead to a solution can however reap rewards in terms of a better allocation of resources. For example, viewing the resource allocation from school districts to individual schools as a Markov Decision Process (MDP) (Howard 2002) can lead to allocations which are more aligned to other aspirations of the educational system (such as the ‘No Child Left Behind’ policy) than allocation methods currently in use (Dimitrov et al. 2014). In this application, a school district comprises a set of schools each of which has a proficiency level (measured by student performance on standardized tests) and the school district aims to maximize aggregate proficiency across schools using its funding allocation model. Each school is modelled as a MDP with states which represent the school’s proficiency level, actions which relate to the school’s funding level, and rewards which emanate from the resources received as a consequence of the proficiency level. The transition probabilities for switching from one proficiency level to another are partly random and partly related to previous funding levels. As such, the probabilities are unknown and so a transition set (of possible values) is used instead. The resulting MDP problem is non-linear and non-convex, but a simplification leads to a tractable problem (Dimitrov et al. 2014). This type of model has the advantage that various possible objectives of the funding bodies, such as equity in distribution of

---

4 Further discussion of using preference functions in the context of university decision making, and also of the need to allow preference functions to evolve within the decision model as stakeholders gain a better understanding of the relevant trade-offs and restrictions, can be found in Dickmeyer (1983).
educational proficiency across schools (BenDavid-Hadar and Ziderman 2011), could be incorporated. This recent work provides an interesting approach to educational resource allocation, and further research into solving such complex problems would be welcome.

In this section we have largely been concerned with planning using forecasts of student numbers, and with the allocation of resources, from the government to the different education sectors, from the sector to the institutions within it, and from the institutions to the departments of which it is comprised. We end with reference to studies which address related but slightly different problems, and which depart from (exclusive) use of traditional OR approaches to planning and resource allocation, such as Markov chains, OCT, and goal programming.

In one study, a management support system is developed to deal with the thorny issue of the allocation of time between the four main activities of university faculty namely supervision, teaching, administration and research (Finlay and Gregory 1994). The aim of the system is to move away from an ‘equal shares across tasks’ approach to one which combines information on individuals’ strengths and weaknesses with the department’s requirements in order to produce an equal total workload (although not necessarily equal task time) across individuals. Whilst the system is an attempt to provide an objective allocation to achieve greater efficiency, the authors recognise the importance of the occasional use of subjective judgement. Given the importance of an individual’s research activity in academic career progression, it is reassuring that the system appears, in practice, to permit changing emphasis from one year to another in terms of time allocated to tasks.

Research performance has become a prominent tool in allocating funding to universities, and this point is recognised in a later study which examines the facilitation of change at organisation level (Nicholls et al. 2004). A mixed-mode modelling methodology combines ‘hard’ OR techniques (for example, social judgement theory used in attitudinal benchmarking, and integer linear programming in identifying research targets) with ‘soft’ approaches (such as change management tools for shaping culture and behaviour). When applied in an Australian university, this methodology results in increased research activity and output (Nicholls et al. 2004). Given the increasing importance of research performance in the allocation of funding to higher education institutions, the relationship between budget allocation and performance is a context where operational researchers can make useful contributions.

3. Efficiency and performance measurement

Planning and resource allocation models are largely concerned with effectiveness i.e. how can a system (best) achieve desired outcomes? In relating inputs to outputs, these models ignore the possibility that production processes may incorporate inefficiencies. A considerable body of OR and related literature is devoted to the measurement of efficiency in education production. Education provides a particularly interesting context for efficiency evaluation because its institutions are both not-for-profit, making conventional measures of performance (such as financial ratios) inappropriate (Berkner 1966), and at least in part publicly funded, leading to public interest in obtaining value for money.

While early work in the OR literature on performance measurement exists (see, for example, Jauch and Glueck 1975), funding cuts in times of austerity have been a major stimulus for studies of
education performance. The 1980s, for example, saw a burgeoning literature on university performance, where deterministic OLS methods were applied to derive value-added measures of teaching output (Johnes and Taylor 1987b; 1987a; 1989a; 1989b; 1990), while research performance was measured using publications counts (Jauch and Glueck 1975). The work on school efficiency, however, eschewed the derivation of school rankings from OLS analyses of aggregate school data. Instead, methods, such as multilevel modelling (MLM) were used. MLM utilises pupil level data in order to isolate measures of school performance which are standardised for pupil achievement on entry to the system and other contextual variables (Aitken and Longford 1986; Woodhouse and Goldstein 1988; Sammons et al. 1993; Goldstein and Spiegelhalter 1996; Goldstein and Thomas 1996).

Education institutions, however, can be seen to be multi-product organizations. Institutions of higher education, for example, produce (in simple terms) teaching, research and third mission (the last reflecting universities’ engagement with society). Schools also produce multiple outputs in terms of education in various subject areas, diverse levels and different types (for example, vocational and academic). Measuring performance of a multi-product organization by separately examining production of each output gives rise to problems, one of which is the interpretation of information over several indicators. Multiple-criteria decision analysis is a field of OR which offers tools for application in this context.

Information can be synthesised using, for example, principal components analysis (Johnes 1996), the analytic hierarchy process (Holder 1990; Tadisina et al. 1991; Holder 1998), co-plot (Paucar-Caceres and Thorpe 2005); or a multi-criteria evaluation, such as a portfolio model where weights are based on preferences (Politis and Siskos 2004; Jessop 2010). These approaches are not without their problems. Principal components analysis, for example, can result in unacceptable loss of information if the first two components used to represent the information do not account for a large percentage of information. The analytic hierarchy process has the potential problem of rank reversal which can occur with the introduction to (or the removal from) the data set of an institution. This violates the principle of rationality, generally considered desirable in any decision analysis technique, that rank order of observations A and B in a given data set should not be affected by the score of observation C (Holder 1998). The co-plot approach, used in multi-criteria decision-making, seeks to display in two dimensions the location of observations on the basis of several attributes simultaneously. It has serious limitations particularly if its underlying assumptions are not satisfied (Mar Molinero and Mingers 2007). Finally a portfolio model requires knowledge of the preferences of those interested in the evaluation.

Another problem with measuring performance using separate indicators relating to each output is that the approach ignores potential synergies in the production of education. Early work which explores jointness in the production of multiple education outputs employs canonical regression analysis (Chizmar and McCarney 1984; Chizmar and Zak 1984; Gyiham-Brempong and Gyapong 1991). While the approach can offer insights into the production process, it does not provide measures of efficiency. Data envelopment analysis (DEA) (Charnes et al. 1978; 1979; Banker et al. 1984) is a non-parametric frontier estimation technique which can handle a production situation.

---

5 For more information see Raveh (2000).

6 A method of deriving efficiency measures from canonical correlation functions is provided in Ruggiero (1998), although there are no applications of this approach to efficiency measurement in the education context.
with both multiple outputs and multiple inputs, and does not require a priori specification of a functional form. DEA therefore is an alternative tool in the context of multi-criteria decision analysis (Doyle and Green 1993); it permits more satisfactory representations of production than the early OLS performance models and is superior to the canonical regression approach in that it provides estimates of efficiency.

Education has been a popular area of application of DEA; indeed the developers of the method demonstrated its application in this context (Charnes et al. 1981); and it is one of the top five areas of application of DEA (Liu et al. 2013). DEA and related non-parametric methods continue to be used to derive measures of efficiency in all sectors of education, including: kindergartens and primary schools (Mancebón and Mar Molinero 2000; Burney et al. 2013); secondary schools (Bradley et al. 2001; Ramanathan 2001; Waldo 2007; Mancebón and Muñiz 2008; Portela and Camanho 2010; Haelermans and De Witte 2012; Haelermans et al. 2012; Haelermans et al. 2012; Haelermans and Ruggiero 2013; Essid et al. 2014; Podinovski et al. 2014); education administrative areas (Smith and Mayston 1987; Ray 1991; Thanassouli and Dunstan 1994; Ruggiero et al. 1995; Ruggiero 1996; Grosskopf et al. 1999; Ruggiero 1999b; 2000; Grosskopf et al. 2001; Grosskopf and Moutray 2001; Fukuyama and Weber 2002; Rassouli-Currier 2007; Ruggiero 2007; Ouellette and Vierstraete 2010); post-compulsory but pre-higher education (Bradley et al. 2010; Johnes et al. 2012); and universities (Beasley 1990; 1995; Mar Molinero 1996; Giménez and Martínez 2006; Fandel 2007; Dehnokhalaji et al. 2010; Thanassouli et al. 2011; Duh et al. 2012; Bayraktar et al. 2013; De Witte et al. 2013). DEA has been used to assess efficiency of individual academic departments or programmes within an institution (Colbert et al. 2000; Kao and Liu 2000; Moreno and Tadepalli 2002; Casu et al. 2005; Kao and Hung 2008; Ray and Jeon 2008), central administration or services across universities (Casu and Thanassouli 2006; Simon et al. 2011), and to make efficiency comparisons across countries of different education systems (Giménez et al. 2007), although care should be taken in interpreting the results of the last given that DEA requires all production units to be comparable and to share a common production environment.

Some particularly novel applications of DEA to education might be noted. First, DEA has been used to evaluate the performance of individuals (rather than organisations), such as pupils in a school (Thanassouli 1999). Application of DEA to individuals within an organization can be extended, however, to produce measures of efficiency of the organization itself, such as a school or university (Portela and Thanassouli 2001; Thanassouli and Portela 2002; Johnes 2006b; Cherchye et al. 2010). Like MLM, this approach allows the effects of individual effort and institutional influence to be disentangled, in this case using a type of meta-frontier approach.\footnote{A related study relaxes the underlying DEA assumption of convexity of the production frontier, allows for uncertainty in the data, and applies the non-parametric free-disposal hull approach to pupil-level data to produce estimates of school efficiency (De Witte et al. 2010). The results of this approach are compared with the results of applying multi-level modelling. The approaches are found to be complementary in the information they provide to managers.}

A second example applies DEA in the context of schools which are ‘differentially effective’ with pupils of different ability. The hypothesis is that schools may have different success (or be differentially effective) in raising the attainment of pupils with low compared to high ability. DEA is developed to identify whether or not such differential effectiveness exists, and to identify peers for any institution looking to alter the direction of their differential effectiveness (Thanassouli 1996a;
The method can be used to identify exemplar schools for those institutions which want to alter the bias of their differential effectiveness.

Third, the possibility of time lags between inputs into and outputs from the production process (such as observed with research activity and dissemination of research findings) are explored in a dynamic model of university production (Emrouznejad and Thanassoulis 2005). Fourth, a network DEA (Kao 2014) is applied in the context of universities each of which is seen as a system in which departments are producing in parallel (Rayeni and Saljooghi 2010). The decomposition of the university production process into its component parts and application of network DEA is further explored by (Johnes 2013). Lastly, DEA is used as a decision-support device to synthesize information produced for different stakeholders into a usable format for each interested party. Examples in the context of applicants to higher education and to university planning are provided, respectively, in Sarrico et al. (1997) and Sarrico and Dyson (2000).

As a deterministic non-parametric approach, DEA has the drawback that there are no conventional tests of significance or methods for drawing inference, and efficiency estimates can be affected by sample size. Particular care should therefore be taken in choosing the inputs and outputs of any DEA model, and the specification should be consistent with the production process being evaluated (Cook et al. 2014). In addition, bootstrapping can be applied to produce bias-corrected estimates of efficiency (see, for example, Essid et al. 2010 in the context of schools), and hypothesis tests have been developed to assess the significance of specific inputs and/or outputs (see, for example, Johnes 2006a in the context of universities). Second stage analyses of the determinants of the DEA efficiency scores abound in the education context; the more recent examples apply bootstrapping procedures to the second stage (Cordero-Ferrera et al. 2010), and to both the first and second stages (Alexander et al. 2010).

A parallel development in frontier estimation methodology can be found in the econometrics literature (Lovell 1995) which saw the introduction of stochastic frontier analysis (SFA) (Aigner et al. 1977; Battese and Corra 1977; Meeusen and van den Broeck 1977). SFA is a parametric frontier estimation method, which, as its name implies, allows for stochastic errors in the data. Following a development by Jondrow et al. (1982) SFA allows the estimation of technical efficiency for each unit of observation. The derivation of efficiency scores combined with estimation of parameters (and hence the potential calculation of economies of scale and scope) means that SFA is particularly popular with education economists. Thus, while numerous applications of SFA to education can be found in the economics and education economics literature (Stevens 2005; McMillan and Chan 2006; Johnes et al. 2008; Abbott and Doucouliagos 2009; Johnes and Schwarzenberger 2011; Kirjavainen 2012; Zoghbi et al. 2013; Johnes 2014) applications are more rare in the OR and management science literature. One explanation is the view that DEA, with its facility to provide benchmarks, is important to managerial decision-making, and hence of relevance to OR. SFA, on the other hand, is used in more policy-oriented applications and hence is perhaps of greater relevance to (education) economists (Lewin and Lovell 1990; Lovell 1995). This is to underestimate the versatility of the techniques, however; each approach can offer both policy and managerial insights (see, for example, Thanassoulis et al. 2011), and both are surely of equal interest to operational
researchers and economists. The scarcity of SFA applications to education in the OR literature is therefore puzzling.

4. Routing and Scheduling

Applications of OR to routing and scheduling problems in the education context are numerous. Examples include the routing and scheduling of buses, the scheduling of courses and examinations, the assignment of pupils to classes and the allocation of students to work groups. The main areas of application are discussed in turn.

4.1 Vehicle routing and scheduling

The transportation of pupils to school is both a routing problem and, since there are time constraints in terms of school start and finish times, a scheduling problem. Whilst the objective of most routing and scheduling problems is to minimize the cost of providing the service (where costs in this context include vehicle depreciation, fuel and drivers), the school bus routing and scheduling problem must take into account other considerations. These considerations, which are of differing levels of concern to different stakeholders, include: the time spent by pupils en route to and from school; time spent waiting at bus stops; overcrowding on buses; and problems for school staff caused by early arrival at and late departure from schools. These problems are invariably set up in an optimization framework. Small-scale problems can be solved using, for example, integer programming methods (Ward 1964; Bektas and Elmaș 2007); more often, though, the complexity of these multi-objective leads to use of heuristics in finding a solution (Bodin and Berman 1979; Schultz 1979; Gochenour Jr. et al. 1980; Corberán et al. 2002; Li and Fu 2002; Spada et al. 2005).

Similar approaches can be applied to the transportation of pupils with special educational needs for whom the gains in terms of reducing the duration of journeys can be particularly beneficial (Russell and Morrel 1986; Sutcliffe and Boardman 1990).

A similar vehicle routing and scheduling problem is encountered in the delivery of school meals. Constraints include: the earliest time for collecting the meal from the kitchens where the meals are cooked; the earliest and latest time of delivery (to be in line with the school timetable) at the dining centre; the period the meal is in transit; the vehicle capacity; and constraints relating to the vehicle drivers. There is evidence that a heuristic approach to solving these complex and highly constrained problems can be simple to use and reduce costs (Atkinson 1990).

Integer programming techniques have been applied in a related problem area: the optimal allocation of contracts. Instances include the allocation of contracts to providers of school meals (Epstein et al. 2002; Epstein et al. 2004) and to bus companies for the transportation of school pupils (Letchford 1996). These examples of optimal assignment represent a category of problem which has received extensive attention in the OR literature.

---

8 The accuracy of the SFA decomposition of the total error into inefficiency and a random component has been questioned, and SFA appears to perform similarly to parametric deterministic models when applied to cross section data (Ruggiero 1999a; Ondrich and Ruggiero 2001). This might partially explain the reluctance of operational researchers to adopt SFA in education applications.

9 The school bus routing problem is a universal one and examples in this section refer to the USA, and countries in Western and Eastern Europe as well as the Far East.

10 This is a context which is likely to be encountered in only a few countries. In England, however, where free school meals for infants were introduced in September 2014, this issue is likely to be of increasing importance.
4.2 The assignment problem

There are many situations where ‘elements’ from a population must be assigned to groups, rooms, institutions or time slots. This is an area in which OR tools can be applied with considerable success.

a) Students to groups

The necessity of assigning students to groups is an issue frequently encountered in the day-to-day running of schools and universities. Typically these types of problems involve creating subsets from a population such that the differences between elements within each subset is maximally diverse (Fan et al. 2011). This maximally diverse grouping is a set-partitioning problem (O’Brien and Mingers 1997). Applications include the assignment of students to classes – for example, forms in schools, seminar groups in universities, or project work groups at any education level – and the allocation of students to university accommodation.

The assignment of students to maximally diverse work groups is of increasing importance particularly in business schools where the focus is to encourage team skills in a diverse setting (Weitz and Lakshminarayanan 1998; Gallego et al. 2013). In one study, integer programming identifies a set of teams which are academically similar, but where the composition of each team is both functionally diverse and demographically balanced – for example, there is no solitary female and no single international student (Cutshall et al. 2007). The programming approach produces teams which both meet the desired criteria and lead to fewer complaints (from students and instructors) compared to previously-used methods. In another example where dissertation students are allocated to seminar streams, goal programming produces groups which meet the preferences of both the students and the seminar organisers (Miyaji et al. 1987); when compared with manual assignment to seminars, goal programming results in a more favourable value of the objective function11.

In creating an assessment and assigning students to groups to undertake that assessment it is not necessarily the students’ (or even the organisers’) ex ante preferences which are important. Rather the assignment of students should be undertaken to create student groups with skill sets that fulfil the requirements of the project (Muller 1989), including achieving the module learning outcomes.

The student assignment problem is sufficiently complex that an integer programming model is not practical because it cannot deliver a solution in an acceptable time limit (Dhar et al. 1990; Weitz and Jelassi 1992). It is therefore common practice to resort to a heuristic approach12 which will identify a solution in an acceptable time frame; the solution will not be an optimal one but can be guaranteed to be of a certain minimum level of quality. Examples include: creating diverse coursework groups in a university (Reeves and Hickman 1992; Weitz and Jelassi 1992); allocating students to their university accommodation (O’Brien and Mingers 1997); and forming large but balanced tutor groups in a comprehensive school (Baker and Benn 2001). Advances in heuristics have allowed increasing

11 In addition to creating student groups, the goal programming approach has also been successfully used to help construct university committees for dealing with promotions where various rules must be applied in the construction (Ceylan et al. 1994), and to assign university faculty to their teaching taking into account preferences for modules and times (Schniederjans and Kim 1987; Ozdemir and Gasimov 2004; Al-Yakoob and Sherali 2006).

12 The efficiency of a number of heuristics is tested and compared in the context of creating student groups (Weitz and Lakshminarayanan 1998).
use of continuous as well as binary characteristics (on which to base the allocation), and the creation of unequal groups. For example, in the case of school tutor groups, the heuristic approach allows groups to be created which are balanced in terms of gender, ability, ethnicity, previous school background and special educational needs. The need for unequal groups is particularly relevant in the context of university accommodation where the size of group is dependent on capacity of shared kitchen which might vary by residence block.

b) Students to education institutions

Similar programming and heuristics methods can be applied to the allocation of pupils to schools. Much of the early literature is concerned with producing an allocation which minimizes total students’ miles (or time) travelled subject to school and class size constraints, and ethnic composition targets (Clarke and Surkis 1967; Belford and Ratliff 1972; McKeown and Workman 1976; Bovet 1982; Schoepfle and Church 1989). The list of desirable targets can become so complex that goal programming might be a more appropriate tool: in addition to pupil travel time and racial balance, schools may need to consider capacity utilization as well as ability and gender balance amongst the pupils (Lee and Moore 1977; Sutcliffe et al. 1984). Goal programming can lead to fewer variables and constraints than linear programming, and can incorporate multiple objectives simultaneously (Knutson et al. 1980).

Some studies highlight the link between allocating pupils to schools and higher level school planning (considered in section 2). Programming methods and decision support systems developed for allocating pupils to schools can indicate those schools which require capacity to be expanded (through additional building), when new schools are required, and the general location of new schools within a district (Gac et al. 2009). In the context of declining numbers, they can also identify which schools should close (Holloway et al. 1975; McKeown and Workman 1976; Jennergren 1980; Ferland and Guénette 1990).13

c) Examination scheduling and course timetabling

Educational timetabling encompasses many variants including the timetabling of courses (together with students, teaching staff and rooms) and examinations (with students and venues). Examination and course scheduling give rise to similar problems, although there are some differences between the two (de Werra 1985; Carter 1986; Qu and Burke 2009).14 For example, several examinations can take place in one room, or one examination could be split across several venues; a class for a course, on the other hand, must take place in a single room (Burke and Petrovic 2002). The increasing flexibility offered to students in terms of number and choices of modules raises the importance of course and examination timetabling in schools and universities: an effective examination or course timetabling decision support system can therefore create huge benefits and time savings. The literature devoted to educational timetabling (examination and course) is vast and spans disciplines. The focus here is on studies reported in OR journals. Surveys of examination timetabling can be found in Carter (1986); Carter et al. (1996); Qu et al. (2009).

13 The role of operational researchers in contributing to decisions and processes surrounding school closure is examined in considerable detail in Mar Molinero (1988; 1993).
14 A related scheduling problem in the education context includes academic conference scheduling. The distinctions between this problem and the examinations and courses timetabling problems are explored in Sampson (2004) and an application can be found in (Eglese and Rand 1987).
Timetabling can be accomplished using one of 7 possible approaches: sequential methods; clustering-based methods; constraint-based methods; meta-heuristic methods; hyper-heuristics; case-based reasoning approaches; and multi-criteria approaches (Carter 1986; Burke and Petrovic 2002; Abdullah et al. 2007).

Sequential methods represent timetabling as a graph-colouring problem\(^{15}\). This somewhat simplistic approach is unlikely to represent the examination scheduling problem satisfactorily. In addition to the need to avoid examination conflicts for students, schedulers also face such hard constraints as: a limited time period over which examinations are taken; size of examination venues; resource supplies determined by examination requirements including, for example, computer facilities; sequential scheduling for certain examinations. There may well also be soft constraints such as: spreading examinations for individual students; scheduling examinations of different lengths in the same location\(^{16}\). In scheduling courses, hard constraints might be that two classes cannot meet in the same room at the same time, and that no class can be spread across more than one room (Carter and Tovey 1992). Soft constraints, particularly in a school setting, might relate to the length of time spent in movement between classes (Hinchcliffe 1973); in a university context students and lecturers might be required to have at least (or at most) a given number of days with contact hours (Burke et al. 2012b). It is imperative that examinations and courses are scheduled such that the hard constraints are met, while it is desirable that the schedule also adheres to the soft constraints. An additional requirement might be that any violations of the soft constraints should be distributed evenly amongst the stakeholders to produce a ‘fair’ timetable (Mühlenthaler and Wanka 2014).

In a sense the scheduling problem is one of achieving feasibility (or a satisficing problem) rather than one of optimization since there is no obvious objective function to be optimized (Johnson 1993)\(^{17}\). The sequential approach therefore assigns events (examinations or lessons) one by one, starting with those considered the most difficult to schedule, attempting to avoid violating the hard constraints at each stage. Various heuristics have been developed to solve the scheduling problem (Petrovic et al. 2007), and there are many examples of their application to course and examination scheduling (Romero 1982; Balakrishnan 1991; Hertz 1998; Carter and Johnson 2001; Burke and Newall 2004; Burke et al. 2010c).

Sequential methods are easy to use and computationally inexpensive; but the quality of their solutions can be mixed, and they really need to be used with additional methods to improve their performance in practice (Petrovic et al. 2007; Mumford 2010). Examples of timetabling which combine graph colouring with meta-heuristics (see below) include: Dowsland (1990); Dowsland and Thompson (1990); Dowsland and Thompson (2005); Burke et al. (2012c).

Clustering-based (or decomposition) methods break the large timetabling problem into smaller sub-problems (Qu et al. 2009) which can be solved relatively easily. Events are split into sub-groups

---

\(^{15}\) The link between graph colouring and timetabling is attributed to Welsh and Powell 1967. Overviews can be found in (de Werra 1985; Carter 1986; de Werra 1997).

\(^{16}\) See Qu et al. (2009) for a comprehensive list of hard and soft constraints which have been used in the literature.

\(^{17}\) In fact, there are a few studies which address timetabling problems using integer or linear programming to optimize an objective function (Gosselin and Truchon 1986; Birbas et al. 1997; Dimopoulou and Miliotis 2004). But the programming framework can lead to problems of such complexity that solution is by a column generation approach (White 1975; Papoutsis et al. 2003; Santos et al. 2012).
which satisfy the hard constraints; sub-groups are then assigned to ‘slots’ to satisfy the soft constraints (Burke and Petrovic 2002). The disadvantage is that the sub-groups are formed and fixed at the start which may result in an unsatisfactory timetable.

**Constraint-based methods** comprise two main approaches: constraint logic programming and constraint satisfaction problem techniques (Qu et al. 2009; Mumford 2010). Examinations are modelled as a set of discrete variables which can take a finite set of values (such as times and rooms). A set of (hard) constraints is also specified and constraint programming is applied to search for an optimal solution (Brailsford et al. 1999). An application to course timetabling can be found in Deris et al. (1997). These methods are generally computationally demanding, and, like sequential methods, are most successful when used in combination with other approaches.

**Meta-heuristic approaches** are useful for finding solutions to complex optimization problems. Examples of meta-heuristics used in timetabling include: simulated annealing, genetic and memetic algorithms (including variations such as harmony search algorithm), greedy randomized adaptive search procedure (GRASP), threshold accepting, great deluge algorithm and tabu search (Petrovic et al. 2007; Lara-Velázquez et al. 2011). These techniques can take account of large numbers of both hard and soft constraints, so they generally provide more satisfactory solutions than other methods, although performance can vary from one problem to another (Petrovic et al. 2007); and this leads to the downside which is that meta-heuristics can be too problem-specific to be adapted cheaply to other problems (Burke et al. 2003a; Burke et al. 2003b). In any case, the operators of timetabling software such as school and university administrators, are unlikely to have the expertise to adapt any meta-heuristic approach to their own specific problem (Pais and Amaral 2012). Lewis (2007) and Pillay (2014b) provide useful surveys of methods in this category, and there are numerous examples of applications in both the course and examinations contexts.

**Multi-criteria approaches** visualise the examination timetabling problem as having a range of optimisation criteria rather than a single one. In a single criterion approach, the weighted costs of violations of different (soft) constraints are used to measure the quality of possible solutions. In practice, different constraints are of varying importance to different parties, but sums of costs fail to reflect this. The multi-criteria approach overcomes this problem by treating each constraint as a criterion to which a specific level of importance is assigned. Thus this approach provides insights into the timetabling problem and offers a degree of flexibility which is not provided by other methods (Burke and Petrovic 2002). Applications to courses and examinations timetabling include: Wood and Whitaker (1998); Burke et al. (2000); Huédé et al. (2006); Beyrouthy et al. (2009).

---

18 Examples of applications to examination scheduling include Lofti and Cerveny (1991); Balakrishnan et al. (1992); Carter et al. (1994); Carter et al. (1996); Carter and Johnson (2001).
19 A meta-heuristic is often developed in the context of a particular problem (or particular class of problem) and its performance outside of this context can therefore be variable.
20 Applications of meta-heuristics to course timetabling include: Barham and Westwood (1978); Tripathy (1980); Abramson (1991); Hertz (1991); Costa (1994); Alvarez-Valdes et al. (1996); Wright (1996); Abramson et al. (1999); Dimopoulou and Miliotis (2001); Mirrazavi et al. (2003); Aladag et al. (2009); Beligiannis et al. (2009); De Causmaecker et al. (2009); Moura and Scaraficci (2010); Zhang et al. (2010); Al-Betar and Khader (2012); Burke et al. (2012b); Geiger (2012); Pais and Amaral (2012); da Fonseca et al. (2014); Lewis and Thompson (2015).
21 Applications of meta-heuristics to examinations timetabling include: Johnson (1990); Thompson and Dowsland (1998); Dimopoulou and Miliotis (2001); White et al. (2004); Abdullah et al. (2007); Burke et al. (2010a); Öçcan et al. (2010); Turabieh and Abdullah (2011); Al-Betar et al. (2014).
Case-based reasoning (CBR) uses past experience to solve new problems (Petrovic et al. 2007). All previously solved examples (timetables in this case) are saved in the computer memory and a new problem is solved by using solutions to similar past problems. CBR therefore generates a new solution to a new problem using an old case which is most similar to the new context. The new problem and new solution are then retained in the computer memory for future use. CBR methods have been shown to outperform graph heuristic methods (in terms of costs of constraint violations), but their success depends on the number and complexity of cases stored (Burke and Petrovic 2002). Applications of CBR can be found in both the course and examinations timetabling context (Burke et al. 2006a; Petrovic et al. 2007).

Hyper-heuristics have their roots in the field of artificial intelligence. They are search methods for choosing, combining or adapting simpler heuristics (or components of heuristics) to address complex problems and have been growing in popularity in the context of examination and lesson timetabling (Qu and Burke 2009; Qu et al. 2009; Pillay 2012) where continuous educational reform and numerous constraints unique to each problem make for highly intricate cases. By operating on a search space of heuristics rather than a search space of solutions, hyper-heuristics have the potential to find the appropriate algorithm (rather than solution) for a specific problem (Burke et al. 2013) and so have much more general applicability than meta-heuristics (Petrovic et al. 2007). The interested reader should consult Pillay (2014a) for a comprehensive survey of the hyper-heuristics literature in the context of educational timetabling. The remainder of this sub-section focuses on the OR literature.

Hyper-heuristics offer considerable potential for application in the context of timetabling and so we briefly consider some different categories. Burke et al. (2010b) propose a classification of hyper-heuristics along two dimensions as illustrated in figure 1. One dimension defines the source of feedback during learning, while the other describes the nature of the heuristic search space. The source of feedback depends on whether hyper-heuristics are non-learning (they do not learn from feedback from the search process) or learning. Non-learning hyper-heuristics are likely to be the least useful in the timetabling context. Learning hyper-heuristics can be online, where the learning happens while the heuristic is solving an instance of the problem, or offline, in which case the learning is from a set of trials and leads to an approach which will generalise to unseen problems. Evidence suggests that algorithms with online learning outperform those which learn offline (Soria-Alcaraz et al. 2014).

The second and orthogonal dimension is the nature of the search space and can be divided into (i) heuristic selection: automated methodologies for choosing existing heuristics, and (ii) heuristic generation: automated methodologies for generating new heuristics from components of existing heuristics. Whilst the ‘methodologies’ are often heuristics, they need not be; indeed an approach such as CBR (discussed previously) has been effectively used in this way in the timetabling context (see Burke et al. 2006b for an example in the context of university course and examination timetabling).

The search space dimension can be further divided into a second level which distinguishes between the classes of lower level heuristics used in the hyper-heuristic framework. These can be described
as construction or perturbation heuristics\textsuperscript{22}. Construction heuristics, as their name implies, construct or build up to a solution from an empty solution set. Thus the hyper-heuristic framework is supplied with a set of construction heuristics and the goal is to select and apply the heuristic which is most appropriate at each decision point.

The graph-colouring problem lends itself to solution by selection construction hyper-heuristics; indeed, the effectiveness of such hyper-heuristics has been demonstrated in the context of both examination and course timetabling (Burke et al. 2007; Qu and Burke 2009). A detailed examination of the characteristics of the heuristic search space of selection construction hyper-heuristics demonstrates why these are effective in solving educational timetabling problems: essentially the landscape of the heuristic search space is globally convex indicating that an optimal solution is not isolated but surrounded by many local minima (Ochoa et al. 2009). Generation construction hyper-heuristics have also been shown to perform satisfactorily in terms of quality of solution and speed and ease of operation in the educational timetabbling setting (Burke and Newall 2004; Pillay and Banzhaf 2009).

In contrast to construction heuristics, perturbation heuristics start with a complete solution and try iteratively to improve on that solution. In this case the hyper-heuristic framework is supplied with a set of neighbourhood structures or simple local searchers. These are iteratively selected and applied to the current solution until a predetermined stopping condition is reached (see, for example, Burke et al. 2003b). The effectiveness of selection perturbation hyper-heuristics in the context of examination and course timetabling has been examined in a number of studies (Moscato and Cotta 2003; Burke et al. 2012a; Burke et al. 2014; Kheiri et al. 2014; Soria-Alcaraz et al. 2014). Research into generation perturbation hyper-heuristics as applied to education timetabbling problems is difficult to find in the OR literature and this is a gap which could usefully be filled in future work.

Timetabling is an issue faced by education institutions around the world, and yet there is still a gap between the theory of timetabbling and practical applications by education administrators. This is partly because the education systems are continuously undergoing reform (and so the requirements of practical applications are constantly changing), and also because each empirical problem has a different and complex set of constraints. Hyper-heuristics aim to provide a general approach which can be applied to any problem, but as yet they are not in common use in practice. Data-sharing might speed up the transition from research into hyper-heuristics into their practical use. For example, using as a base the different timetabbling systems faced in five countries – Australia, England, Finland, Greece and the Netherlands – Post et al. (2012) develop and share an Extensible Markup Language (XML) format of the general timetabbling problem. The facilitation of such data exchange between researchers and practitioners in the field could lead to a narrowing of the gap currently observed between the theory and practice of timetabbling (McCollum et al. 2012; McCollum and Burke 2014).

5. Conclusion

This paper provides an overview of diverse areas in education where OR tools have been applied. It shows education as a field in which operational researchers have been, and continue to be, active.

\textsuperscript{22} We therefore have 4 types of hyper-heuristics along this dimension: selection construction; selection perturbation; generation construction; and generation perturbation.
and successful in providing improved operations and solutions. It is therefore surprising that, with special issues of OR journals devoted to applications in various fields, none has been devoted to OR applications in education.

OR techniques have offered insights at the macro-level into the allocation of resources to the education sector and the division of these between the different levels of education and the institutions within the sectors. A considerable amount of effort has been devoted to the efficiency of educational institutions, and here developments in multi-criteria decision analysis and frontier estimations methods have proved useful in synthesizing information across indicators and in providing efficiency measures and benchmarks for institutions to improve their performance. Scheduling and timetabling are also areas where OR has been able to contribute to the education field. Bus routing and scheduling and examination and lesson timetabling are more efficient as a consequence.

The OR in education literature still has many gaps where operational researchers could make useful contributions. While primary, secondary and tertiary education are well researched, there is little on non-compulsory (i.e. pre-higher) or vocational education. E-learning also goes largely unnoticed despite the recent surge in massive open online courses (MOOCs). It is suggested that the growth in virtual learning could cause a massive restructuring of education provision, particularly in higher education (Scott 2001), and operational researchers are well placed to advise on the best organisational structure of the higher education sector in these times of change. E-learning also offers advantages in terms of data generation; educational data mining is a new research area which might be used in conjunction with e-learning in developing and designing courses (Romero and Ventura 2007).

Hyper-heuristic approaches are increasingly being researched in course and examination timetabling applications (Burke et al. 2007; Pillay and Banzhaf 2009; Qu and Burke 2009; Pillay 2012), and the general nature of their approach means that they are likely to play a key role in increasingly complex timetabling problems faced by educational administrators. Much of the work has been in the context of university course and examination timetabling, and so future work needs to include school timetabling as well (Pillay 2014a). The greatest contribution in this area would be the development of easy-to-use software which would produce satisfactory solutions for application in any practical timetabling context.

Given the complexities of real world educational systems, performance measurement is an area which continues to present many opportunities for research (Mayston 2003), and OR has an obvious role to play in the ongoing audit and development of performance indicators in education (Dyson 2000). Despite a considerable OR literature devoted to performance measurement, there have been few attempts to evaluate the costs of inefficiency in education. The one exception suggests that the losses from inefficiency in secondary education are substantial (Taylor 1994). There is therefore clearly scope for more work in this context.

---

23 Some exceptions include: Kodama (1970); Qing and Mar Molinero (1988); Abbott and Doucouliagos (2002); Bradley et al. (2010); Johnes et al. (2012).

24 With the exception of an early paper by Jamison and Lumsden (1975).

25 This point is emphasised by Peter Drucker: ‘Thirty years from now the big university campuses will be relics. Universities won't survive. It's as large a change as when we first got the printed book.’ (Lenzner and Johnson 1997).
Given that enhanced efficiency is a universal goal (and that efficiency measurement is necessary to achieve that goal) recent developments in DEA, for example, dynamic DEA and network DEA, could prove useful in the education context. In addition, new developments in SFA, such as random parameter and latent class modelling, might tempt operational researchers to apply SFA to education.

While frontier estimation methods determine the education production frontier (and hence allow us to evaluate whether we are ‘doing things right’), it is surely of interest to know that we are simultaneously ‘doing the right things’. This leads to a need to assess both the efficiency and effectiveness of education institutions (Powell et al. 2012). An early attempt at this (Golany 1988) demonstrates that operational researchers have the tools to apply to this problem which involves the integration of frontier estimation methods with multi-objective methods to assess effectiveness in achieving educational goals.

Over more than 50 years OR has contributed much to the operations of the education sector. Rapid changes in the way in which education can be delivered, together with the pressing need to provide education more efficiently ensure that there will continue to be a need for operational researchers to address and offer solutions to the problems observed in the education context.
Figure 1: A classification of hyper-heuristics proposed by Burke et al. (2010b)

<table>
<thead>
<tr>
<th>Feedback</th>
<th>Nature of the heuristic research space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online learning</td>
<td>Heuristic selection</td>
</tr>
<tr>
<td>Offline learning</td>
<td></td>
</tr>
<tr>
<td>No learning</td>
<td>Construction heuristics</td>
</tr>
</tbody>
</table>

**Heuristic selection**
- Methodologies to select

**Heuristic generation**
- Methodologies to generate

References


